ChatGPT 的科学计量分析

岳磊[*]
资阳环境科技职业学院 资阳 641300

摘要：[目的/意义] ChatGPT 作为新一代人工智能，对社会生活各方面均可能产生影响，梳理 ChatGPT 科学研究的现状，为相关研究的开展提供参考。[方法/过程] 通过科学计量的分析方法，对知网包含的 161 篇 ChatGPT 文献进行了分析。[结果/结论] ChatGPT 目前的研究分布在 40 个不同的领域，学者间合作还不够紧密，其研究主要集中在沿海地区。在 ChatGPT 的研究中，主要涉及到对于 ChatGPT 技术的探讨，ChatGPT 对于各行业的影响，尤其是其对教育的影响。当前 ChatGPT 的研究还处于早期阶段，需要更多的实证研究去验证其在现实生活中的实际影响。

关键词：ChatGPT；科学计量；R；VOSviewer

1 引言

2022 年 11 月 30 日，OpenAI 发布 ChatGPT，短短五天该应用程序的注册用户便突破百万。仅两个月后，其更是达到了惊人的上亿用户。一石激起千层浪，Meta、百度、谷歌、阿里等国内外互联网头部企业都争相研发相关产品。与此同时，各行各业开始加入到有关于 ChatGPT 的讨论之中，ChatGPT 在网络媒体上的热度也持续升温（见图 1）。

ChatGPT（Chat Generative Pre-trained Transformer），中文称为“生成型预训练变换模型”，它是一款人工智能技术驱动的自然语言处理工具，该工具可以完成脚本撰写、代码编写、翻译、对话等任务。相较于传统的人工智能，ChatGPT 有着更大的数据来源、更强的算力和算法以及高度的人智化等特征[1]。一方面，ChatGPT 可以提高工作效率、优化服务等等[1-3]。另一方面，ChatGPT 的使用也有许多的风险和伦理需要进行考虑[4, 5]。

作为爆款人工智能，ChatGPT 在国内也被广泛关注，有部分期刊已经刊登了有关于 ChatGPT 的论文。本研究基于中国知网探究 ChtaGPT 的研究现状，为认识和研究 ChatGPT 提供资讯。

作者邮箱：yniannian911@outlook.com

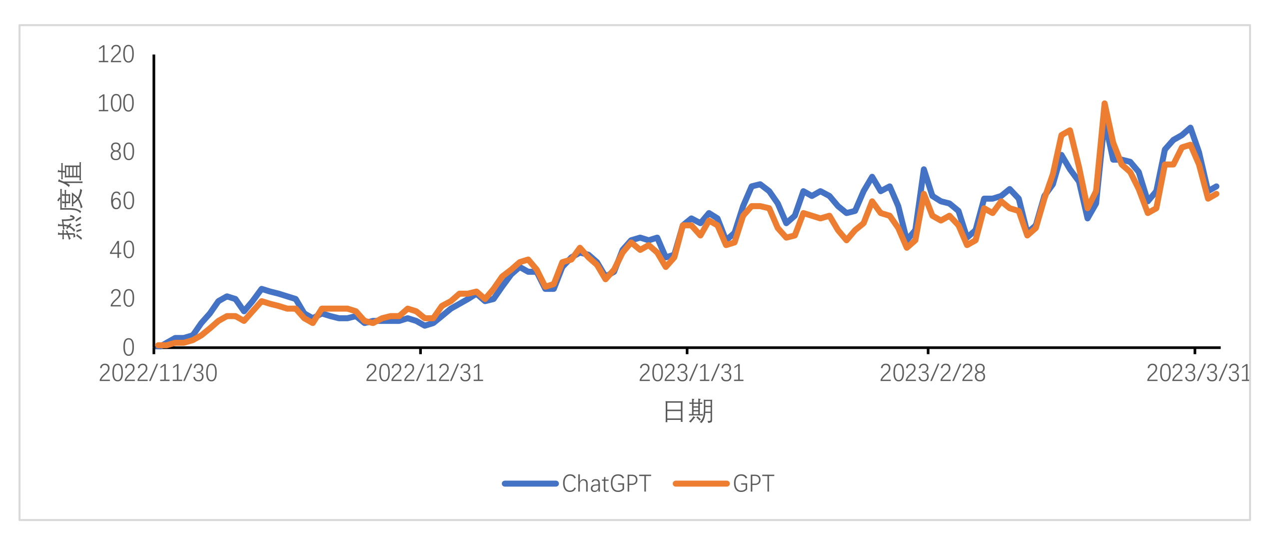


图 1 2022 年 11 月 30 日至 2023 年 4 月 2 日 ChatGPT 的谷歌趋势

注：这些数字代表搜索词在特定区域和时间范围内的热门程度变化趋势，以图表中的最高点做为比较基准。100 分代表该字词的热门程度在该时间点达到最高峰，50 分表示该字词的热门程度为最高点的一半，0 分则表示该字词热门程度的资料不足。

2 数据与方法

使用知网为数据来源，以“ChatGPT”为检索词进行主题检索，检索日期为 2023 年 4 月 3 日。结果显示，报纸为 448 篇，学术期刊为 381 篇（其中中文库为 161 篇）。

本研究使用科学计量的方法对中文库的 161 篇学术期刊文献的数据进行统计分析，并采用 R、VOSviewer 等软件对结果进行可视化[6-8]。本研究中，对收集到的 161 篇 ChatGPT 中文库文献进行两个方面的分析：（1）ChatGPT 文献产出的描述性分析；（2）ChatGPT 的研究热点。

3 研究结果

3.1 ChatGPT 文献的描述性统计

在领域维度上，根据知网数据，ChatGPT 的研究分布在 40 个不同的领域。其中论文产出量排在前列的领域有自动化技术（103 篇）、信息经济与邮政经济（36 篇）、计算机软件及计算机应用（33 篇）、教育理论与教育管理（25 篇）、新闻与传媒（12 篇）、图书情报与数字图书馆（10 篇）。ChatGPT 作为人工智能的新一代领军产品，可以处理各种语言任务。因此，目前自动化技术是其主要聚焦的领域。

ChatGPT 研究的地区分布如图 2A 和 2B，当前我国有 19 个地区发表了关于 ChatGPT 的

论文。整体上发表论文的区域集中在沿海省份。北京以 77 篇名列第一，第二和第三分别是上海（24 篇）和广东（10 篇）。

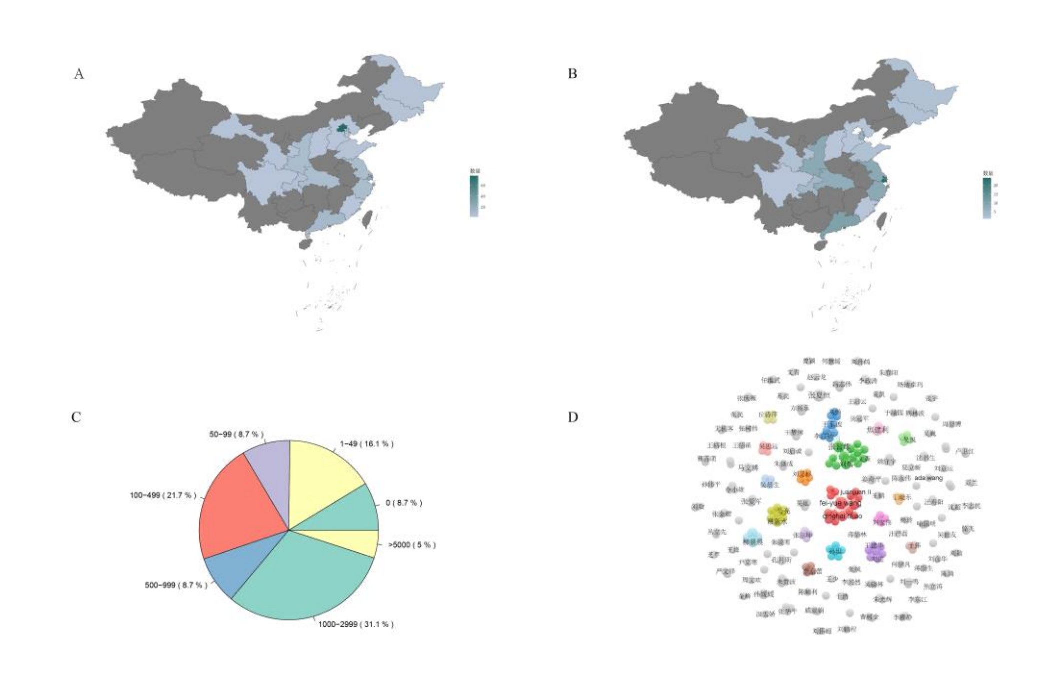


图 2 ChatGPT 研究的描述性统计分布.（A）ChatGPT 研究的空间分布（含北京）;（B）ChatGPT 研究的空间分布（不含北京）;（C）ChatGPT 研究的下载频次分布;（D）ChatGPT 研究的作者共现图谱。

在作者合作上（如图 2D），有 65 篇文章的作者数量为 2 个及其以上，占总文章数的 40.34%。结合图 2D 可以看出，整体上合作网络呈现高度的不连通状态，ChatGPT 作为新生事物，科研人员在合作强度上稍显不足，未来在该领域存在着很强的合作空间。

下载量是论文在期刊网站、出版商或数据库中被保存、下载的次数，是出版数字化下的产物[9]。对于新兴事物（如本研究的 ChatGPT）的研究来讲，下载量相对于被引量可能更为灵敏。分析 161 篇 ChatGPT 文献的下载频次分布，筛选出高下载量论文。结果显示（图 2C），有超过 1/3 的论文下载量高于 1000，ChatGPT 在中国科研人员中也备受关注。

3.2 ChatGPT 研究的热点主题分析

对 ChatGPT 的关键词进行分析[1]，结果见图 3。一共形成 3 大聚类： 1）ChatGPT 的技术;（3）ChatGPT 对于教育的影响;（2）ChatGPT 对于其他行业的影响。

聚类 1 是关于 ChatGPT 技术的探讨。钱力等人梳理了 ChatGPT 技术的整体架构图，并

[1] 本次图谱参数设置如下：关键词最小共现次数为 2，将 gpt 替换为 chatgpt，忽视图谱中未连通的关键词，最小聚类尺度为 4。

系统分析和讨论了 ChatGPT 的语料体系、预训练算法与模型和微调算法与模型[10]。郭全中和张金熠总结了 ChatGPT 的四大技术特征（单一模态、巨量数据、人工标注以及强化学习）[11]，朱光辉和王喜文对其运行模式和关键技术也进行了分析[12]。

聚类 2 是 ChatGPT 对于教育领域的影响。吴晓林和邢羿飞认为 ChatGPT 作为新一代人工智能可能会冲击以知识传递为中心的教育模式，但其本身也存在局限，在短时间内不太可能取代人类教学[13]。钟秉林等人通过笔谈分析了 ChatGPT 热潮下其应用对于高等教育的挑战以及驱动教育变革[14]。沈超也分析了 ChatGPT 作为新技术如何驱动教育变革以及促进创新人才的培养[15]。焦建利认为 ChatGPT 可以助推学校教育的数字化转型[16]。王佑镁等人则辩证地分析了 ChatGPT 的应用潜能和风险[17]。

聚类 3 是 ChatGPT 对于除教育领域外其他行业的影响。张夏恒分析了 ChatGPT 对于数字经济发展的影响，ChatGPT 可以促进教育、医疗、传媒、电商等多领域的数字化转型，提升各领域的数字化效率，拓宽各领域数字化边界，提升数字化经济高质量发展[18]。周智博讨论了 ChatGPT 引入数字政府建设时的风险以及如何对 ChatGPT 的应用进行规制[19]。社会工作者、会计、情报工作等多个领域也有研究者进行了讨论[20-22]。

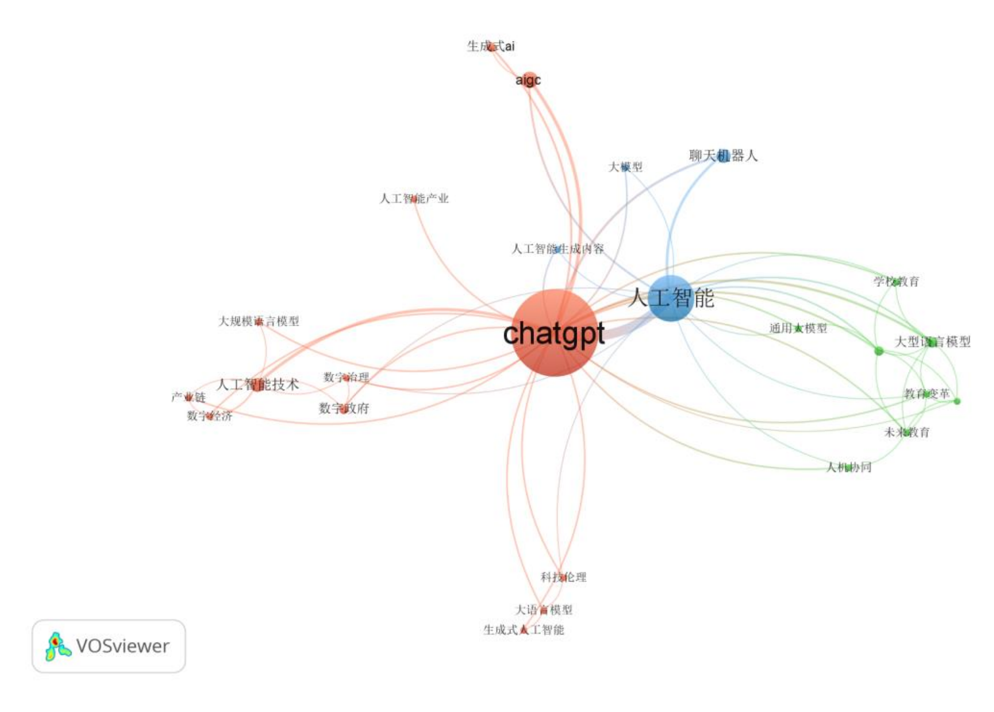


图 3 ChatGPT 的关键词图谱

4 讨论

本文从科学计量的角度，系统分析了目前国内关于 ChatGPT 的研究现状，以期为未来

ChatGPT 的研究提供信息参考。本文得出以下结论：

（1）以中国知网检索的结果来看，ChatGPT 作为热点领域，整体发文规模目前并不活跃，但仍有新的论文不断产出。从空间分布上看，文献主要分布在沿海地区，这与经济发展的地区差异有一定的一致性。从下载量来看，总体上 ChatGPT 作为社会热点被广泛关注。

（2）在研究主题上，ChatGPT 的研究主要涉及技术和应用的讨论。在领域分布上，ChatGPT 的中文研究涉及到 40 个不同的领域，各行业的研究者和从业者都开始注意到 ChatGPT 对于所在领域的冲击或影响。

综上，ChatGPT 作为新生事物其研究还处于早期阶段，当前的研究主要集中在技术上的讨论以及从理论角度去探讨其应用的影响，还缺乏大量的实证研究去探讨 ChatGPT 对在实践中对于各行业以及个体各方面的真实影响。ChatGPT 作为能力卓越的爆款人工智能，禁止和逃避已不可行。因此，研究者在未来仍需用实际数据说话，去探讨类 ChatGPT 人工智能对社会生活各方面的影响，以此来让公众辩证看待类 ChatGPT 人工智能的优劣。

参考文献

[1] 郑世林，姚守宇，王春峰. ChatGPT 新一代人工智能技术发展的经济和社会影响[J/OL]. 产业经济评论(2023-03-12). https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD65x2y26DsGNMTH8yl6tCbOmm4rmvEgJzN3JlzBpx0SHWJTS0z6aJFKb&uniplatform=NZKPT

[2] van Dis E A M, Bollen J, Zuidema W, et al. ChatGPT: five priorities for research[J]. Nature, 2013, 614: 224-226.

[3] 张夏恒. 新一代人工智能技术（ChatGPT）及其对人类社会的影响与变革[J/OL]. 产业经济评论(2023-03-14). https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD65x2y26DsGNMTH8yl6tCbOmolJJr0wUvftVV3oNteyXI1h1mb3i4SKf&uniplatform=NZKPT

[4] 令小雄，王鼎民，袁健. ChatGPT 爆火后关于科技伦理及学术伦理的冷思考[J]. 新疆师范大学学报(哲学社会科学版), 2023, 44(04): 123-136.

[5] Gao C A, Howardet F M, Markovet N S, et al. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers[EB/OL]. https://doi.org/10.1101/2022.12.23.521610

[6] 李杰. 科学知识图谱原理及应用[M]. 北京：高等教育出版社, 2018.

[7] R Core Team. R: A language and environment for statistical computing [computer software]. 2022, R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

[8] Van Eck N J, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping [J]. Scientometrics, 2010, 84(2): 523-38.

[9] 谢娟, 龚凯乐, 成颖, 等. 论文下载量与被引量相关关系的元分析[J]. 情报学报, 2017, 36(12): 1255-1269.
[10] 钱力, 刘熠, 张智雄, 等. ChatGPT 的技术基础分析[J/OL]. 数据分析与知识发现(2023-03-26). https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD66BsvSnCfrndXMzn9cFE5goLTRzDvPVZLb-IMaKhtD1BpWawVYd6jkT&uniplatform=NZKPT
[11] 郭全中, 张金熠. ChatGPT 的技术特征与应用前景[J]. 中国传媒科技, 2023, (01): 159-160.
[12] 朱光辉, 王喜文. ChatGPT 的运行模式、关键技术及未来图景[J]. 新疆师范大学学报(哲学社会科学版), 2023. 44, (04): 113-122.
[13] 吴晓林, 邢羿飞. 知识复制抑或创新激发？——人工智能(ChatGPT)对社科类研究生教育的挑战与机遇[J]. 广西师范大学学报(哲学社会科学版)(2023-03-22).
[14] 钟秉林, 尚俊杰, 王建华, 等. ChatGPT 对教育的挑战（笔谈）[J/OL]. 重庆高教研究(2023-03-02). https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD64rqXKKfkFrrKh6io4uIsstkLFG-qpqDiuWYnZMDDPAXo1QRCFhxOBE&uniplatform=NZKPT
[15] 沈超. ChatGPT：助力高等教育变革与创新型人才培养[J]. 国家教育行政学院学报, 2023, (03): 13-16.
[16] 焦建利. ChatGPT 助推学校教育数字化转型——人工智能时代学什么与怎么教[J]. 中国远程教育, 2023, 43(04): 16-23.
[17] 王佑镁, 王旦, 梁炜怡, 等. “阿拉丁神灯”还是“潘多拉魔盒”：ChatGPT 教育应用的潜能与风险[J]. 现代远程教育研究, 2023, 35(02): 48-56.
[18] 张夏恒. 基于新一代人工智能技术(ChatGPT)的数字经济发展研究[J/OL]. 长安大学学报(社会科学版)(2023-02-23). https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD64MAtUjjuPFifTiF93u0FBZImVkSNNZW3LzfAvAESnRmEFqhXaWr9Ws&uniplatform=NZKPT
[19] 周智博. ChatGPT 模型引入我国数字政府建设：功能、风险及其规制[J/OL]. 山东大学学报(哲学社会科学版)(2023-03-15). https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD66FeCiybgjTV5edK5ZLn9yQHHDEDz98C5WCpzcXjcUq9VK3C5l0jBuf&uniplatform=NZKPT
[20] 吴世友. ChatGPT 会取代社会工作者吗[J]. 中国社会工作, 2023, (06): 25-26.
[21] 刘勤. ChatGPT 及其对会计工作的影响探讨[J]. 会计之友, 2023, (06): 158-161.
[22] 张智雄, 于改红, 刘熠, 等. ChatGPT 对文献情报工作的影响[J/OL]. 数据分析与知识发现(2023-04-01). https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD66BsvSnCfrndXMzn9cFE5god6CXruF4eWR1iUBbQRzInPD4l65is5ws&uniplatform=NZKPT

Scientometrics Analysis of ChatGPT Literatures

Lei Yue

Ziyang College of Environmental Science and Technology

Abstract: [Objective/Significance] ChatGPT, as a new generation of artificial intelligence, may have an impact on all aspects of social life. This paper reviews the status quo of scientific research on ChatGPT and provides reference for future related research. [Method/Process] Using scientometrics analysis to analyze 161 articles searched from CNKI. [Results/Conclusion] ChatGPT's current research is distributed in 40 different fields. The cooperation among scholars is not close enough, and its research is mainly concentrated in coastal areas. The research on ChatGPT mainly involves the discussion on ChatGPT technology and its influence on various industries, especially its influence on education. At present, ChatGPT's research is still at an early stage, and more empirical studies are needed to verify its practical impact in real life.

Keywords: ChatGPT; Scientific metrology; R; VOSviewer